

# Privacy-preserving Average Consensus: Privacy Analysis and Optimal Algorithm Design

Jianping He<sup>1</sup>, Lin Cai<sup>1</sup>, Chengcheng Zhao<sup>2</sup>, and Peng Cheng<sup>2</sup>

**Abstract**—The goal of the privacy-preserving average consensus (PPAC) is to guarantee the privacy of initial states and asymptotic consensus on the exact average of the initial value. This goal is achieved by an existing PPAC algorithm by adding and subtracting variance decaying and zero-sum random noises to the consensus process. However, there is lack of theoretical analysis to quantify the degree of the privacy protection. In this paper, we analyze the privacy of the PPAC algorithm in the sense of the maximum disclosure probability that the other nodes can infer one node's initial state within a given small interval. We first introduce a privacy definition, named  $(\epsilon, \delta)$ -data-privacy, to depict the maximum disclosure probability. We prove that PPAC provides  $(\epsilon, \delta)$ -data-privacy, and obtain the closed-form expression of the relationship between  $\epsilon$  and  $\delta$ . We also prove that the added noise with uniform distribution is optimal in terms of achieving the highest  $(\epsilon, \delta)$ -data-privacy. Then, we prove that the disclosure probability will converge to one when all information used in the consensus process is available, i.e., the privacy is compromised. Finally, we propose an optimal privacy-preserving average consensus (OPAC) algorithm to achieve the highest  $(\epsilon, \delta)$ -data-privacy. Simulations are conducted to verify the results.

**Index Terms**—Average consensus, Data privacy, Optimal algorithm, Disclosure probability.

## I. INTRODUCTION

Consensus has attracted extensive attention over the past decades, since it is an efficient algorithm for distributed computing and control. A consensus algorithm refers to the action that nodes in the network reach a global agreement regarding a certain opinion using their local neighbors' information only [1]. Due to the robustness and scalability, consensus has been applied in a variety of areas, e.g., coordination and cooperation [2], [3], distributed estimation and optimization [4], [5], sensor fusion [6], distributed energy management [7] and scheduling [8], and time synchronization in sensor networks [9]–[11].

Average consensus is the most commonly adopted consensus algorithm, where the agreement reached by the algorithm equals the average of all nodes' initial values. For traditional average consensus algorithms, each node will broadcast its real state value to neighbor nodes during consensus process. Hence, under traditional average consensus algorithms, the state information of each node is disclosed to its neighbor nodes. However, in some applications, the initial values of nodes are private information, which means that nodes do not want to release their real initial values to other nodes

[16]. For example, consensus algorithm is adopted in social networks for a group of members to compute the common opinion on a subject [18]. In this application, each member may want to keep his personal opinion on the subject secret to other members. Also, considering the multi-agent rendezvous problem [19], a group of nodes want to eventually rendezvous at a certain location, while the participators may not want to release their initial locations to others. This means that when the privacy is concerned, each node's real state value may not be available to the other nodes, and thus the traditional consensus algorithm becomes invalid.

Recently, researchers have investigated the privacy-preserving average consensus problem, which aims to guarantee that the privacy of initial state is preserved while average consensus can still be achieved [12]–[17]. The basic idea is to add random noises to the real state value during the communication to protect the privacy, and then carefully design the noise adding process such that average consensus is achieved. For example, Huang et al. [13] designed a differentially private iterative synchronous consensus algorithm by adding independent and exponentially decaying Laplacian noises to the consensus process. Their algorithm can guarantee differential privacy. As the algorithm may converge to a random value, the exact average consensus may not be guaranteed. Nozari et al. [14] pointed out and proved that it is impossible to achieve average consensus and differential privacy simultaneously. Hence, they design a novel linear Laplacian-based consensus algorithm, which guarantees that an unbiased estimate of the average consensus can be achieved almost surely with differential privacy. Manitara and Hadjicostis [15] proposed a privacy preserving average consensus algorithm by adding correlated noises to the consensus process. The proposed algorithm guarantees the initial state of each node cannot be perfectly inferred by the other “malicious” nodes. A sufficient condition is provided under which the privacy of benign agents' initial states are preserved. More recently, Mo and Murray in [16] well addressed the privacy-preserving average consensus problem by designing a novel Privacy Preserving Average Consensus (PPAC) algorithm, where exponentially decaying and zero-sum normal noises are added to the consensus process. They proved that the PPAC algorithm achieves the exact average consensus in the mean-square sense, and also proved in their extension work [17] that the algorithm achieves minimum privacy breach in the sense of disclosed space.

However, there is lack of theoretical results to quantify the degree of the privacy protection and what is the relationship between estimation accuracy and privacy. To fill this gap, in this paper, we provide theoretical privacy analysis for the

1: The Dept. of Electrical & Computer Engineering at the University of Victoria, BC, Canada jphe@uvic.ca; jianpinghe.zju@gmail.com; cai@ece.uvic.ca

2: The State Key Lab. of Industrial Control Technology and Innovation Joint Research Center for Industrial Cyber Physical Systems, Zhejiang University, China zccsq90@gmail.com; pcheng@iipc.zju.edu.cn

existing PPAC algorithm [16] in the sense of the maximum disclosure probability that other nodes can infer one node's initial state within a given small interval (a given estimation accuracy). A privacy definition, named  $(\epsilon, \delta)$ -data-privacy, which is first introduced in our previous work [22], is exploited to depict the maximum disclosure probability. This privacy definition reveals the relationship between the privacy and the estimation accuracy. We provide theoretical results to quantify the degree of the privacy preservation and demonstrate the quantitative relationship of the estimation accuracy and the privacy under the PPAC algorithm. Based on the analysis, it is found that the uniform distribution noise is the optimal one in terms of achieving the highest  $(\epsilon, \delta)$ -data-privacy, and the exact initial state of a node can be perfectly inferred, i.e., privacy is compromised, when a node has all information used in the consensus process. Hence, to solve this problem, we design a novel OPAC algorithm to achieve the exact average consensus as well as data-privacy. The main contributions of this paper are summarized as follows.

- We prove that the existing PPAC algorithm provides  $(\epsilon, \delta)$ -data-privacy, and obtain a closed-form expression of the relationship between the estimation accuracy and the privacy (the relationship between  $\epsilon$  and  $\delta$ ).
- We prove that for the added random noises, the uniform distribution is optimal in the sense that PPAC algorithm can achieve the highest privacy when the mean and variance of noises are fixed.
- We prove that when all the information used in the consensus process are available for estimation, the privacy of the initial state will exponentially decay with iteration and eventually be lost. This result reveals how the exact initial state can be inferred, and is consistent with the privacy analysis given in [17].
- We design an OPAC algorithm to achieve average consensus while guarantee the highest  $(\epsilon, \delta)$ -data-privacy. We prove that OPAC algorithm converges to the exact average and provides the highest  $(\epsilon, \delta)$ -data-privacy even if all the information used in the consensus process are available for estimation.

The remainder of this paper is organized as follows. Section II provides the introduction of consensus and the important definitions. Section III introduces the PPAC algorithm and formulates the problem. In Section IV, we provide theoretical results on the degree of privacy protection. The OPAC algorithm is proposed in Section V. Section VI verifies the main results and conclusions are given in Section VII.

## II. PRELIMINARIES

The network is abstracted as an undirected and connected graph,  $G = (V, E)$ , where  $V$  is the set of nodes and  $E$  is the set of the communication links (edges) between nodes.  $(i, j) \in E$  if and only if (iff) nodes  $i$  and  $j$  can communicate with each other. Let  $N_i$  be the neighbor set of node  $i$ , where  $j \in N_i$  iff  $(i, j) \in E$ , i.e.,  $N_i = \{j | j \in V, (i, j) \in E, j \neq i\}$ .

### A. Average Consensus

Suppose that there are  $n$  ( $n \geq 3$ ) nodes in the network (i.e.,  $|V| = n$ ), and each node  $i$  has an initial scalar state  $x_i(0)$ . For

an average consensus algorithm, each node will communicate with its neighbor nodes and update its state based on the received information to obtain the average of all initial state values. Hence, the traditional average consensus algorithm is given as follows,

$$x_i(k+1) = w_{ii}x_i(k) + \sum_{j \in N_i} w_{ij}x_j(k), \quad (1)$$

for  $\forall i \in V$ , which can be written in the matrix form as

$$x(k+1) = Wx(k), \quad (2)$$

where  $w_{ii}$  and  $w_{ij}$  are weights, and  $W$  is the weight matrix. It is well known from [20] that if, 1)  $w_{ii} > 0$  and  $w_{ij} > 0$ ; and 2)  $W$  is doubly stochastic matrix, then an average can be achieved by (1), i.e.,

$$\lim_{k \rightarrow \infty} x_i(k) = \frac{\sum_{\ell=1}^n x_\ell(0)}{n} = \bar{x} \quad (3)$$

When the privacy of nodes' initial states are concerned, all nodes are unwilling to release its real state to the neighbor nodes at each iteration. It means that each  $x_j(k)$  is unavailable in (1). To preserve the privacy of nodes' initial states, a widely used approach is to add a random noise to the real state value when a node needs to communicate with its neighbor nodes at each iteration. We define a new state as

$$x_i^+(k) = x_i(k) + \theta_i(k), i \in V, \quad (4)$$

where  $\theta_i(k)$  is the added random noise for privacy preservation at iteration  $k$ . With the adding noise process, the update equation (1) is changed to,

$$x_i(k+1) = w_{ii}x_i^+(k) + \sum_{j \in N_i} w_{ij}x_j^+(k) \quad (5)$$

$$= w_{ii}[x_i(k) + \theta_i(k)] + \sum_{j \in N_i} w_{ij}[x_j(k) + \theta_j(k)], \quad (6)$$

for  $\forall i \in V$ . Therefore, a privacy-preserving average consensus algorithm is to design the added noises (including the distribution and the correlations among them), such that the goal of (1) is achieved under (5).

### B. Privacy Definitions

We develop the following two privacy definitions which will be used later in the paper. By referring to [22], we first give the definition of  $(\epsilon, \delta)$ -data-privacy as follows.

*Definition 2.1:* A distributed algorithm provides  $(\epsilon, \delta)$ -data-privacy, if the probability that each node  $i$  can successfully estimate its neighbor node  $j$ 's initial value  $x_j(0)$  in a given interval  $[x_j(0) - \epsilon, x_j(0) + \epsilon]$  is no larger than  $\delta$ , i.e.,

$$\delta = \max_{\hat{x}_j(0)} \Pr\{\hat{x}_j(0) \in [x_j(0) - \epsilon, x_j(0) + \epsilon]\}, \quad (7)$$

where  $\hat{x}_j(0)$  is one of the estimation of  $x_j(0)$ .

Under (4) and (5), each node  $j$  broadcasts  $x_j^+(k)$  to its neighbor nodes at each iteration  $k$ , and then its neighbor node  $i$  can infer/estimate the initial value based on the received information,  $\{x_j^+(k) | k = 0, 1, \dots, \infty\}$ , and its own states. Since the value of each broadcast information  $x_j^+(k)$  equals the value of real state plus a noise, node  $i$  will take the

probability over the space of all noises  $\{\theta_j(k)|k = 0, 1, \dots, \infty\}$  to estimate the values of the added noises, and then get the estimation of  $\hat{x}_j(0)$ .

**Definition 2.2:** Suppose that two algorithms  $A_1$  and  $A_2$  provide  $(\epsilon, \delta_1)$ -data-privacy and  $(\epsilon, \delta_2)$ -data-privacy for a given  $\epsilon$ , respectively. Then, we say that  $A_1$  achieves a higher  $(\epsilon, \delta)$ -data-privacy than  $A_2$ , if,  $\delta_1 < \delta_2$ .

The  $(\epsilon, \delta)$ -data-privacy is different from the differential privacy. It is well known that differential privacy aims to provide means to maximize the accuracy of queries from statistical databases while minimizing the chances of identifying its records [21]. However,  $(\epsilon, \delta)$ -data-privacy aims to minimize the maximum disclosure probability  $\delta$  under any given estimation accuracy  $\epsilon$ .

### III. PROBLEM FORMULATION

The PPAC-algorithm proposed to guarantee the privacy of initial state and the exact average consensus by adding zero-sum and exponential decaying normal distribution noises to the consensus process of (5) [16]. The following is the brief description of the PPAC algorithm.

- 1) At iteration  $k$ , each node generates a normal distributed random variable  $\nu_i(k)$  with mean 0 and variance  $\sigma^2 = 1$ . Assume that all the random variables  $\nu_i(k), i = 1, \dots, n; k = 0, 1, \dots$  are jointly independent.
- 2) Each node uses a random noise  $\theta_i(k)$  in (4) to get  $x_i^+(k)$ , where

$$\theta_i(k) = \begin{cases} \nu_i(k) & \text{if } k = 0 \\ \varrho^k \nu_i(k) - \varrho^{k-1} \nu_i(k-1) & \text{otherwise} \end{cases} \quad (8)$$

where  $\varrho \in (0, 1)$  is a constant for all nodes.

- 3) Each agent then communicates with its neighbors and updates its state with (5).
- 4) Let  $k = k + 1$  and go to step 1).

The authors in [16] have proved that average consensus can be achieved by their PPAC algorithm in the mean-square sense. In their extension work [17], they also proved that the privacy is compromised when the disclosed space of a node with  $m$  neighbors is of dimension  $m + 1$ . In this paper, we will first investigate the privacy of PPAC algorithm based on the definition of  $(\epsilon, \delta)$ -data-privacy, and then design an optimal privacy-preserving average consensus (OPAC) algorithm in terms of  $(\epsilon, \delta)$ -data-privacy protection. In summary, we will consider the following four critical problems: i) we consider to quantify and analyze the privacy of PPAC algorithm; ii) we study how will the distribution and correlation of the random adding noises affect the privacy; iii) whether and how a node's exactly initial state can be inferred by the other nodes when these nodes can utilize all information used in the consensus process; iv) how to design the OPAC algorithm to achieve the optimal  $(\epsilon, \delta)$ -data-privacy and the exact average consensus.

### IV. PRIVACY ANALYSIS OF PPAC

Before presetting main results, we first give the basic assumptions and the information set used for state estimation. Assume that the distribution and the correlation of the random variable  $\nu_i(k), k = 0, 1, \dots$ , and the update rule of PPAC are

available to all nodes. The full topology information and  $n$  are assumed to be unknown to each node  $i$ , which means that each node cannot know the neighbor set of its neighbor nodes and the total node number of the whole network. The initial states of nodes are assumed to be independent with each other. For estimation, if there is no information of a variable, then the variable could be any value in  $R$ . Then, define two information sets of node  $i$  at iteration  $k$  as follows,

$$\mathcal{I}_i^0(k) = \{x_i^+(0), \dots, x_i^+(k)\}, \quad (9)$$

and

$$\mathcal{I}_i^1(k) = \{N_i, w_{ii}, w_{ij}, x_i^+(0), x_j^+(0), \dots, x_i^+(k), x_j^+(k) | j \in N_i\}. \quad (10)$$

The information set  $\mathcal{I}_i^0(k)$  only includes the states  $x_i^+(\ell), \ell = 0, 1, \dots, k$ , which are used for communication at iteration  $\ell$ . Thus, its neighbor nodes can easily obtain  $\mathcal{I}_i^0(k)$  by storing the information received from node  $i$  at each iteration. The information set  $\mathcal{I}_i^1(k)$  includes all information used in consensus process (6) for node  $i$ . Other nodes can obtain these information by eavesdropping attack.

#### A. Privacy of PPAC Algorithm

In this subsection, based on the definition of  $(\epsilon, \delta)$ -data-privacy, we first analyze the privacy of PPAC algorithm and reveal the relationship between the privacy and the estimation accuracy, under which only  $\mathcal{I}_i^0(k)$  is available. We state a theorem as follows, where its proof method has referred to the proof of Theorem 4.1 in [22].

**Theorem 4.1:** Suppose that the information set  $\mathcal{I}_i^0(k)$  of node  $i$  is available to the other nodes. Then, the PPAC algorithm achieves  $(\epsilon, \delta)$ -data-privacy, where the  $\delta$  satisfies

$$\delta = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{y^2}{\sigma^2}\right) dy \quad (11)$$

and  $\lim_{\epsilon \rightarrow 0} \delta = 0$ .

**Proof:** To prove this theorem, we first prove that at each iteration  $k$ , the probability that each node  $j$  can successfully infer that  $x_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon]$  is no larger than  $\delta$  based on the information set  $\mathcal{I}_i^0(k)$ .

At time  $k = 0$ , node  $j$  can estimate  $x_i(0)$  based on the fact that

$$x_i^+(0) = x_i(0) + \theta_i(0), \quad (12)$$

where  $x_i^+(0)$  and the distribution of  $\theta_i(0)$  are known to  $j$ . Let the corresponding estimation be  $\hat{x}_i(0)$  satisfying

$$\hat{x}_i(0) = x_i^+(0) - \hat{\theta}_i(0). \quad (13)$$

Then, given an estimation  $\hat{x}_i(0)$ , we have

$$\begin{aligned} & \Pr \left\{ \hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^0(0) \right\} \\ &= \Pr \left\{ |\hat{\theta}_i(0) - \theta_i(0)| \leq \epsilon | \mathcal{I}_i^0(0) \right\} \\ &= \Pr \left\{ \theta_i(0) \in [\hat{\theta}_i(0) - \epsilon, \hat{\theta}_i(0) + \epsilon] | \mathcal{I}_i^0(0) \right\} \\ &= \int_{\hat{\theta}_i(0) - \epsilon}^{\hat{\theta}_i(0) + \epsilon} f_{\theta_i(0)}(y) dy, \end{aligned} \quad (14)$$

where  $f_{\theta_i(0)}(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{y^2}{\sigma^2}\right)$  is the probability density function (PDF) of  $\theta_i(0)$ . Since  $\hat{\theta}_j(0)$  is an estimation of  $\theta_i(0)$  and could be any constant, one infers that

$$\begin{aligned} & \Pr\{\hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^0(0)\} \\ & \leq \max_{t \in R} \int_{t-\epsilon}^{t+\epsilon} f_{\theta_i(0)}(y) dy \\ & = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{y^2}{\sigma^2}\right) dy. \end{aligned} \quad (15)$$

At time  $k = 1$ , node  $j$  can estimate  $x_i(0)$  based on  $\mathcal{I}_i^0(1)$ . Thus, it can use the fact of both (12) and the following equation for estimation,

$$\begin{aligned} \frac{x_i^+(1)}{w_{ii}} &= \frac{x_i(1) + \theta_i(1)}{w_{ii}} \\ &= x_i^+(0) + \sum_{l \in N_i} \frac{w_{il}}{w_{ii}} x_l^+(0) + \frac{1}{w_{ii}} \theta_i(1) \\ &= x_i(0) + \theta_i(0) + \frac{1}{w_{ii}} \theta_i(1) + \sum_{l \in N_i} \frac{w_{il}}{w_{ii}} x_l^+(0). \end{aligned} \quad (16)$$

If only using (12), as the proof at iteration  $k = 0$ , (15) holds. Then, we consider the case that node  $j$  uses (16) for estimation. Define

$$\begin{aligned} \theta_i^+(1) &= \theta_i(0) + \frac{1}{w_{ii}} \theta_i(1) + \sum_{l \in N_i} \frac{w_{il}}{w_{ii}} x_l^+(0) \\ &= \theta_i(0) + \theta_i^-(1). \end{aligned} \quad (17)$$

Since the initial states of nodes are independent with each other and the topology information is not available for inferring,  $\sum_{l \in N_i} \frac{w_{il}}{w_{ii}} x_l^+(0)$  in (17) could be any value in  $R$ , thus  $\theta_i^-(1)$  could be any value in  $R$ . Hence, when the value of  $\theta_i^-(1)$  is released, node  $j$  cannot increase the estimation accuracy of  $\theta_i(0)$  with (17) although  $\theta_i^-(1)$  and  $\theta_i(0)$  are coupled (since  $\frac{1}{w_{ii}} \theta_i(1)$  included in  $\theta_i^-(1)$  is coupled with  $\theta_i(0)$ ). Let  $\hat{\theta}_i^+(1)$  be an estimation of  $\theta_i^+(1)$  and  $f_{\theta_i^+(1)}(y)$  be the corresponding PDF of  $\theta_i^+(1)$ . Then, with (16) and (17), we have

$$\begin{aligned} & \Pr\{\hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^+(1)\} \\ & \leq \Pr\{\hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^+(1), w_{ii}, \theta_i^-(1)\} \\ & = \int_{\hat{\theta}_i^+(1)-\epsilon}^{\hat{\theta}_i^+(1)+\epsilon} f_{\theta_i^+(1)|\theta_i^-(1)}(y) dy \\ & \leq \max_{t \in R} \int_{t-\epsilon}^{t+\epsilon} f_{\theta_i(0)}(y) dy \\ & = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{y^2}{\sigma^2}\right) dy. \end{aligned} \quad (18)$$

Meanwhile, note that one node can combine (12) and (16)

together for estimation. In this case, we have

$$\begin{aligned} & \Pr\{\hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^0(1)\} \\ & \leq \max_{t_1, t_2 \in R} \int_{t_1-2\epsilon}^{t_1+2\epsilon} \int_{t_2-2\epsilon}^{t_2+2\epsilon} f_{\theta_i(0), \theta_i^+(1)}(y, z) dz dy \\ & \leq \max_{t_1, t_2 \in R} \int_{t_1-2\epsilon}^{t_1+2\epsilon} \int_{t_2-2\epsilon}^{t_2+2\epsilon} f_{\theta_i^+(1)|\theta_i(0)}(z|y) f_{\theta_i(0)}(y) dz dy \\ & \leq \max_{t \in R} \int_{t-\epsilon}^{t+\epsilon} f_{\theta_i(0)}(y) dy \\ & = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{y^2}{\sigma^2}\right) dy. \end{aligned} \quad (19)$$

From (15), (18) and (19), one concludes that

$$\begin{aligned} & \max_{\hat{x}_i(0)} \Pr\{\hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^0(1)\} \\ & = \max_{\hat{x}_i(0)} \Pr\{\hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^0(0)\}, \end{aligned} \quad (20)$$

which means that the value of  $\delta$  will not increase with iteration, although there are more information of  $\mathcal{I}_i^0(1)$  than  $\mathcal{I}_i^0(0)$ .

Similarly, the result of (20) also holds for each iteration  $k$ , i.e.,

$$\begin{aligned} & \max_{\hat{x}_i(0)} \Pr\{\hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^0(k)\} \\ & = \max_{\hat{x}_i(0)} \Pr\{\hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^0(0)\}. \end{aligned} \quad (21)$$

Hence, one infers that

$$\begin{aligned} & \Pr\{\hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^0(k)\} \\ & \leq \frac{1}{\sigma\sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{y^2}{\sigma^2}\right) dy. \end{aligned} \quad (22)$$

Since the above equation holds for  $\forall k$ , we have (11). Therefore, the PPAC algorithm is  $(\epsilon, \delta)$ -data-privacy. Meanwhile,

$$\lim_{\epsilon \rightarrow 0} \delta = \lim_{\epsilon \rightarrow 0} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{y^2}{\sigma^2}\right) dy = 0.$$

We thus have completed the proof.  $\blacksquare$

From the above proof, it is observed that the privacy is not decreased with iteration when only the information set  $\mathcal{I}_i^0 (= \{\mathcal{I}_i^0(k) | k = 0, 1, \dots\})$  is available for estimation. The main reason is that based on  $\mathcal{I}_i^0$ , node  $j$  cannot know the neighbor set information of node  $i$ , so that after one iteration there are unknown information embedded into the  $x_i^+(k)$  for  $k \geq 1$ . Hence, after one iteration, using  $x_i^+(k)$  for  $k \geq 1$  cannot improve the estimation accuracy. Also, one can see that the value of  $\delta$  does not depend on the estimation approaches. Hence, from (14) and (21), we can state a corollary as follows.

**Corollary 4.2:** Suppose that the added noises used in (6) are jointly independent. If only the information set  $\mathcal{I}_i^0$  is available for node  $j$  to estimate, we have

$$\begin{aligned} & \max_{\hat{x}_i(0), k \in \mathbf{N}^+} \Pr\{\hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^0(k)\} \\ & = \max_{\hat{x}_i(0)} \Pr\{\hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^0(0)\} \end{aligned} \quad (23)$$

$$= \max_{\hat{\theta}_i(0)} \int_{\hat{\theta}_i(0)-\epsilon}^{\hat{\theta}_i(0)+\epsilon} f_{\theta_i(0)}(y) dy, \quad (24)$$



i.e., the relationship between the privacy and the estimation accuracy always satisfies (24), and the privacy is not decreased with iteration.

From Corollary 4.2, one infers that for a given  $\epsilon$ ,  $\delta$  equals to the maximum disclosure probability of  $x_i(0)$  and can be calculated by (24), i.e.,

$$\delta = \max_{\hat{\theta}_i(0)} \int_{\hat{\theta}_i(0)-\epsilon}^{\hat{\theta}_i(0)+\epsilon} f_{\theta_i(0)}(y) dy.$$

Clearly,  $\delta$  only depends on  $f_{\theta_i(0)}(y)$  and  $\epsilon$  since the estimation  $\hat{\theta}_i(0)$  could be any value in the domain of  $\theta_i(0)$ . Thus,  $\delta$  is a function of  $f_{\theta_i(0)}(y)$  and  $\epsilon$ , i.e.,  $\delta = \delta(f_{\theta_i(0)}(y), \epsilon)$ . Based on Definition 2.2, a smaller  $\delta$  can provide a higher  $(\epsilon, \delta)$ -data-privacy for any given  $\epsilon$ . Then, we aim to find the optimal distribution for  $\theta_i(0)$  such that the algorithm can achieve the highest  $(\epsilon, \delta)$ -data-privacy.

*Remark 4.3:* It should be noticed that the above results in this subsection are obtained under the assumption that the topology information is unknown to the nodes and the initial value could be any value in  $R$ . If these assumptions are relaxed, the above results could not be true for the PPAC algorithm in some cases. For example, if the topology information is available and node  $i$  has only one neighbor node  $j$ , then in (17),  $\sum_{l \in N_i} \frac{w_{il}}{w_{ii}} x_l^+(0) = \frac{w_{ij}}{w_{ii}} x_j^+(0)$  is available node  $j$ . Thus, for  $\theta_i^+(1)$ , only

$$\theta_i(0) + \frac{1}{w_{ii}} \theta_i(1) = \nu_i(0) + \frac{1}{w_{ii}} (\rho \nu(1) - \nu(0))$$

is random, which could have a smaller variance than  $\theta_i(1)$  (e.g., when  $w_{ii} = \frac{3}{4}$  and  $\rho = \frac{1}{4}$ ). It means that the value of  $\delta$  will be increased in this case, and thus (11) and (23) cannot be guaranteed. When the topology information is known and all the communication information of node  $i$  and its neighbor nodes is available for estimation, we need to design a new algorithm to guarantee the average consensus and the data privacy simultaneously, which will be discussed in Sec. V.

### B. Optimal Noise Distribution

In this subsection, we find an optimal distribution for the noise adding process in the sense of achieving the highest  $(\epsilon, \delta)$ -data-privacy for the PPAC algorithm. Note that a smaller  $\epsilon$  means a higher accuracy estimation. It means that when  $\epsilon$  becomes smaller, the value of  $\delta$  is more important for the privacy preservation. Hence, we define the optimal distribution for privacy concerns as follows.

*Definition 4.4:* Let  $f_{\theta_i(0)}^*(y)$  be the optimal distribution of  $\theta_i(0)$ , it means that for any given distribution  $f_{\theta_i(0)}^1(y)$ , there exists an  $\epsilon_1$  such that  $\delta(f_{\theta_i(0)}^*(y), \epsilon) < \delta(f_{\theta_i(0)}^1(y), \epsilon)$  holds for  $\forall \epsilon \in (0, \epsilon_1]$ .

To obtain the optimal distribution described in Definition 4.4, we define  $\arg \min_{f_{\theta_i(0)}(y)} \delta = f_{\theta_i(0)}^*(y)$ . Then, we formulate the following minimization problem,

$$\begin{aligned} \min_{f_{\theta_i(0)}(y)} \quad & \delta \\ \text{s.t.} \quad & \mathbf{E}\{\theta_i(0)\} = 0, \\ & \mathbf{Var}\{\theta_i(0)\} = \sigma^2. \end{aligned} \quad (25)$$

The solution of (25) is the optimal distribution for the added noises with given mean and variance used in the PPAC algorithm in terms of  $(\epsilon, \delta)$ -data-privacy.

*Theorem 4.5:* Suppose that the information set  $\mathcal{I}_i^0(k)$  of node  $i$  is available to other nodes. The optimal solution of problem (25) is that

$$f_{\theta_i(0)}^*(y) = \begin{cases} \frac{1}{2\sqrt{3}\sigma} & \text{if } y \in [-\frac{1}{\sqrt{3}\sigma}, \frac{1}{\sqrt{3}\sigma}], \\ 0 & \text{otherwise,} \end{cases} \quad (26)$$

i.e., the uniform distribution is optimal.

*Proof:* We prove this theorem by contradiction. Without loss of generality, we assume that  $\sigma^2 = \frac{1}{3}$ . Let  $f_1(y)$  and  $f_2(y)$  be the PDF of two random variables with mean 0 and variance  $\sigma^2 = \frac{1}{3}$ , and they follow a uniform and non-uniform distribution, respectively. Clearly, we have

$$f_1(y) = \begin{cases} \frac{1}{2}, & \text{if } y \in [-1, 1], \\ 0, & \text{otherwise.} \end{cases} \quad (27)$$

Suppose that the non-uniform distribution  $f_2(y)$  is the optimal distribution. From Definition 4.4, there exists an  $\epsilon_2$ , such that

$$\max_{t \in R} \int_{t-\epsilon}^{t+\epsilon} f_1(y) dy > \max_{t \in R} \int_{t-\epsilon}^{t+\epsilon} f_2(y) dy, \quad (28)$$

holds for  $\forall \epsilon \in (0, \epsilon_2]$ . Since the above equation holds for arbitrarily small value of  $\epsilon$ , we infer that

$$\max_{y \in R} f_1(y) > \max_{y \in R} f_2(y).$$

Since  $f_1(y)$  is a uniform distribution satisfying (27),

$$f_1(y) - f_2(y) > 0, \quad y \in [-1, 1].$$

It directly follows that

$$\int_{-1}^1 f_1(y) dy - \int_{-1}^1 f_2(y) dy > 0. \quad (29)$$

From the definition of a PDF, we have  $\int_{-1}^1 f_1(y) dy = 1$ . Then, we infer from (29) that

$$\int_{-1}^1 f_2(y) dy < 1. \quad (30)$$

Since both  $f_1(y)$  and  $f_2(y)$  have mean 0 and variance  $\sigma^2 = \frac{1}{3}$ , we have

$$\int_{-\infty}^{+\infty} f_1(y) y^2 dy - \int_{-\infty}^{+\infty} f_2(y) y^2 dy = 0, \quad (31)$$

which means that

$$\int_{-1}^1 (f_1(y) - f_2(y)) y^2 dy = \left( \int_{-\infty}^{-1} + \int_1^{+\infty} \right) f_2(y) y^2 dy. \quad (32)$$

For the left hand side of (32), we have

$$\begin{aligned} \int_{-1}^1 (f_1(y) - f_2(y)) y^2 dy &< \int_{-1}^1 (f_1(y) - f_2(y)) dy \\ &= 1 - \int_{-1}^1 f_2(y) dy \end{aligned} \quad (33)$$

For the right hand side of (32), since we have  $\int_{-\infty}^{+\infty} f_2(y)dy = 1$  and (30), it holds that

$$\begin{aligned} \left( \int_{-\infty}^{-1} + \int_1^{+\infty} \right) f_2(y)y^2 dy &> \left( \int_{-\infty}^{-1} + \int_1^{+\infty} \right) f_2(y)dy \\ &= 1 - \int_{-1}^1 f_2(y)dy \end{aligned} \quad (34)$$

Combining (32), (33) and (34) renders a contradiction that

$$\begin{aligned} 1 - \int_{-1}^1 f_2(y)dy &< \int_{-1}^1 (f_1(y) - f_2(y)) y^2 dy \\ &< 1 - \int_{-1}^1 f_2(y)dy. \end{aligned} \quad (35)$$

Hence, we cannot find a non-uniform distribution  $f_2(y)$  such that the value of  $\delta$  is smaller than that under uniform distribution  $f_1(y)$ . It means that the uniform distribution is the optimal solution of (25). Then, based on the definition of uniform distribution, it is not difficult to obtain (26), which completes the proof. ■

Therefore, from the above theorem, one concludes that if we use uniform distribution noises to substitute the normal distribution noises in step 1 of PPAC, it can provide the optimal  $(\epsilon, \delta)$ -data-privacy. Using uniform distribution noises, we have  $\delta = \frac{\epsilon}{\sqrt{3}\sigma}$ , i.e.,

$$\min_{f_{\nu_i}(y)} \max_{\hat{x}_i(0)} \Pr \left\{ \hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^0 \right\} = \frac{\epsilon}{\sqrt{3}\sigma}. \quad (36)$$

Clearly, given a small  $\epsilon (\ll \sigma^2)$ , we have

$$\frac{\epsilon}{\sqrt{3}\sigma} < \frac{1}{\sigma\sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{y^2}{\sigma^2}\right) dy,$$

which means that the privacy of PPAC is enhanced.

### C. Privacy Compromission

In this subsection, we reveal that for PPAC, when  $\mathcal{I}_i^1(k)$  (including more information, e.g., the topology information and information used in consensus process than  $\mathcal{I}_i^0(k)$ ) is available to other nodes for estimation, the exact initial state of node  $i$  can be perfectly inferred, and thus the privacy of the initial state is compromised.

*Theorem 4.6:* Suppose that the information set  $\mathcal{I}_i^1(k)$  of node  $i$  is available to the other nodes. Then, we have

$$\begin{aligned} &\Pr \left\{ \hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^1(k) \right\} \\ &\leq \frac{1}{\varrho^k \sigma \sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{y^2}{\varrho^{2k} \sigma^2}\right) dy, \end{aligned} \quad (37)$$

i.e., the privacy is decaying with iteration, and

$$\lim_{k \rightarrow \infty} \Pr \left\{ \hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^1(k) \right\} = 1, \quad (38)$$

i.e.,  $x_i(0)$  is disclosed and the privacy is compromised.

*Proof:* Based on the information set  $\mathcal{I}_i^1(k)$ , the information of weights and states used in (5) are available. That is the state sequence  $x_i(1), x_i(2), \dots, x_i(k)$  of node  $i$  is released to other nodes. Then, with (4), one obtains the values of

$\theta_i(1), \theta_i(2), \dots, \theta_i(k)$ . Thus, when  $k > 0$ , all the adding noises and the states of node  $i$  are available to other nodes, except  $x_i(0)$  and  $\theta_i(0)$ .

Then, under information set  $\mathcal{I}_i^1(k)$ , one can combine (12), the correlation between  $\theta_i(0)$  and  $\{\theta_i(1), \theta_i(2), \dots, \theta_i(k)\}$  together to estimate the value of  $x_i(0)$ . According to step 2, one obtain the correlation of the added noises, which is given by

$$\begin{aligned} \sum_{\ell=0}^k \theta_i(\ell) &= \theta_i(0) + \sum_{\ell=1}^k [\varrho^\ell \nu_i(\ell) - \varrho^{\ell-1} \nu_i(\ell-1)] \\ &= \nu_i(0) - \varrho^0 \nu_i(0) + \varrho^1 \nu_i(1) - \varrho^1 \nu_i(1) + \varrho^2 \nu_i(2) \\ &\quad - \dots - \varrho^{k-1} \nu_i(k-1) + \varrho^k \nu_i(k) \\ &= \varrho^k \nu_i(k) = \phi_i(k). \end{aligned} \quad (39)$$

Then, adding  $\sum_{\ell=1}^k \theta_i(\ell)$  to both sides of (12) yields that

$$\sum_{\ell=1}^k \theta_i(\ell) + x_i^+(0) = x_i(0) + \sum_{\ell=0}^k \theta_i(\ell) = x_i(0) + \phi_i(k). \quad (40)$$

With (40), the corresponding estimation can be described as

$$\hat{x}_i(0) = \sum_{\ell=1}^k \theta_i(\ell) + x_i^+(0) - \hat{\phi}_i(k). \quad (41)$$

Hence, it follows that

$$\begin{aligned} &\Pr \left\{ \hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^1(k) \right\} \\ &= \Pr \left\{ |\hat{\phi}_i(k) - \phi_i(k)| \leq \epsilon | \mathcal{I}_i^1(k) \right\} \\ &= \Pr \left\{ \phi_i(k) \in [\hat{\phi}_i(k) - \epsilon, \hat{\phi}_i(k) + \epsilon] | \mathcal{I}_i^1(k) \right\} \\ &= \int_{\hat{\phi}_i(k) - \epsilon}^{\hat{\phi}_i(k) + \epsilon} f_{\phi_i(k)}(y) dy. \end{aligned} \quad (42)$$

Note that  $\phi_i(k) = \varrho^k \nu_i(k)$  and  $\nu_i(k) \sim N(0, \sigma^2)$ , and one follows that  $\phi_i(k) \sim N(0, \varrho^{2k} \sigma^2)$ . Then, from (42), we have

$$\begin{aligned} &\Pr \left\{ \hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^1(k) \right\} \\ &= \int_{\hat{\phi}_i(k) - \epsilon}^{\hat{\phi}_i(k) + \epsilon} \frac{1}{\varrho^k \sigma \sqrt{2\pi}} \exp\left(-\frac{y^2}{\varrho^{2k} \sigma^2}\right) dy \\ &\leq \frac{1}{\varrho^k \sigma \sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{y^2}{\varrho^{2k} \sigma^2}\right) dy, \end{aligned} \quad (43)$$

which means that (37) holds.

Meanwhile, we note that  $\lim_{k \rightarrow \infty} \varrho^{2k} \sigma^2 = 0$ , which means that the variance of  $\phi_i(k)$  will converge to 0 when  $k \rightarrow \infty$ . Then, from (39), one infers  $\lim_{k \rightarrow \infty} \sum_{\ell=0}^k \theta_i(\ell) = 0$ . It means

$$\theta_i(0) = - \sum_{\ell=1}^{\infty} \theta_i(\ell). \quad (44)$$

Since  $\theta_i(1), \theta_i(2), \dots, \theta_i(k)$  are available under  $\mathcal{I}_i^1(k)$  for any integer  $k$ ,  $\theta_i(0)$  is inferred with (44) when  $k \rightarrow \infty$ . Hence, both  $x_i^+(0)$  and  $\theta_i(0)$  in (12) are disclosed when  $k \rightarrow \infty$ , i.e.,  $x_i(0)$  is disclosed and (38) holds. ■

Similarly, if  $\nu_i(k)$  follows a uniform distribution, we have

$$\begin{aligned} & \min_{f_{\nu_i}(y)} \max_{\hat{x}_i(0)} \Pr \left\{ \hat{x}_i(0) \in [x_i(0) - \epsilon, x_i(0) + \epsilon] | \mathcal{I}_i^1(k) \right\} \\ &= \min \left\{ 1, \frac{\epsilon}{\sqrt{3} \varrho^k \sigma} \right\}. \end{aligned} \quad (45)$$

Unfortunately, with the information set of  $\mathcal{I}_i^1(k)$ ,  $k = 0, 1, \dots$ , so long as the added noises are jointly independent we cannot decrease the disclosure probability when  $k \rightarrow \infty$ , i.e., the privacy will be disclosed. The reason is that  $\sum_{\ell=0}^{\infty} \theta_i(\ell) = 0$  is a necessary condition to guarantee average consensus when the added noises among nodes are independent (the proof is omitted due to the space limitation), and then (44) is always available for nodes to infer  $\theta_i(0)$ . Hence, the privacy of the initial state is compromised under  $\mathcal{I}_i^1(k)$  when  $k \rightarrow \infty$  even the correlation of the noises is changed.

*Remark 4.7:* The proof of Theorem 4.6 reveals that how the exact initial state of one node can be inferred by the other nodes with the information set of  $\mathcal{I}_i^1(\infty)$  ( $= \{\mathcal{I}_i^1(k) | k = 0, 1, \dots, \infty\}$ ). Theorem 4 in [17] has pointed out that the disclosed space of an node with  $m$  neighbor is of dimension  $m + 1$ . Clearly, the dimension of the information set of  $\mathcal{I}_i^1$  is  $|N_i| + 1$ . Hence, one concludes that Theorem 4.6 is consistent with Theorem 4 in [17].

## V. OPAC ALGORITHM

In this section, we design an OPAC algorithm to achieve the highest  $(\epsilon, \delta)$ -data-privacy, and at the same time to avoid privacy to be compromised even if the information  $\mathcal{I}_i^1(\infty)$  of each node  $i$  is available to other nodes.

### A. Algorithm Design

From the privacy analysis in the above section, we note that the uniform distribution is optimal for the added noise in terms of achieving the highest  $(\epsilon, \delta)$ -data-privacy with  $\delta = \frac{\epsilon}{\sqrt{3}\sigma}$  (given variance  $\sigma$ ). Hence, in each iteration of the OPAC algorithm, we will use uniform distribution noise. We also note that the privacy is compromised when  $\mathcal{I}_i^1(\infty)$  is available. It is because that the nodes can use  $\mathcal{I}_i^1(\infty)$  to obtain the real values of  $\theta_i(1), \theta_i(2), \dots, \theta_i(\infty)$ , and then use the correlation  $\sum_{k=0}^{\infty} \theta_i(k) = 0$  to infer  $\theta_i(0)$ , and thus the value of  $x_i(0)$  is revealed. To avoid the privacy compromise in this case, we introduce a secret continuous function  $F_{ij}(z) : R \rightarrow R$  for node  $i$  with respect to its neighbor node  $j$ . Suppose that  $F_{ij}(z)$  and  $F_{ji}(z)$  are only available to nodes  $i$  and  $j$ . Then, the OPAC algorithm is described as follows.

### B. Convergence and Privacy Analysis

In this subsection, we analyze the convergence and the privacy of the OPAC algorithm.

*Theorem 5.1:* Using the OPAC algorithm, we have (3) holds for  $\forall i \in V$ , i.e., an exact average is achieved.

*Proof:* From Theorem 4.1 of [22], we know that if the added noises in (4) are bounded and decaying, and the sum of all nodes' added noises equals zero, then the average consensus can be achieved. In the following, we prove that

---

### Algorithm 1 : OPAC Algorithm

---

- 1: **Initialization:** Each node  $i$  selects a uniform distribution random variable  $\nu_i(0)$  from interval  $[-\frac{1}{\sqrt{3}\sigma}, \frac{1}{\sqrt{3}\sigma}]$ , and arbitrarily selects a constant sequence  $z_{ij} (\in R)$  for  $j \in N_i$ .
- 2: Let  $\theta_i(0) = \nu_i(0)$  and  $\mathbf{x}_i^+(0) = \mathbf{x}_i(0) + \theta_i(0)$ , then each node  $i$  transmits  $\mathbf{x}_i^+(0)$  and  $z_{ij}$  to its neighbor node  $j$ .
- 3: Each node  $i$  calculates  $\tilde{\nu}_i(0)$  by

$$\tilde{\nu}_i(0) = \nu_i(0) - \sum_{j \in N_i} [F_{ij}(z_{ij}) - F_{ji}(z_{ji})], \forall i \in V. \quad (46)$$

- 4: **Iteration:** Each node updates its state with (5).
- 5: Each node generates a uniform distribution random variable  $\nu_i(k)$  from interval  $[-\frac{1}{\sqrt{3}\sigma}, \frac{1}{\sqrt{3}\sigma}]$  for  $k \geq 1$ .
- 6: Each node  $i$  uses  $\theta_i(k)$  in (4) to get  $x_i^+(k)$ , where

$$\theta_i(k) = \begin{cases} \varrho \nu_i(1) - \tilde{\nu}_i(0) & \text{if } k = 1 \\ \varrho^k \nu_i(k) - \varrho^{k-1} \nu_i(k-1) & \text{if } k \geq 2 \end{cases} \quad (47)$$

where  $\varrho \in (0, 1)$  is a constant for all nodes.

- 7: Each node  $i$  communicates with its neighbors with  $x_i^+(k)$ .
  - 8: Let  $k = k + 1$  and go to step 4.
- 

the added noises used for the OPAC algorithm satisfy these conditions.

We first prove that the added noises are bounded and exponentially decaying. Clearly,  $\theta_i(0) = \nu_i(0) \in [-\frac{1}{\sqrt{3}\sigma}, \frac{1}{\sqrt{3}\sigma}]$  is bounded. Since each  $F_{ij}(z)$  is continuous function, its value is bounded for any given  $z$ . Then, it follows from (46) that  $\tilde{\nu}_i(0)$  is bounded. For  $k \geq 1$ , because  $\nu_i(k)$  is selected from interval  $[-\frac{1}{\sqrt{3}\sigma}, \frac{1}{\sqrt{3}\sigma}]$  and  $\theta_i(k)$  is generated by (47), it is not difficult to infer that each  $\theta_i(k)$  is bounded. Meanwhile, it follows from (47) that

$$\begin{aligned} \lim_{k \rightarrow \infty} |\theta_i(k)| &\leq \lim_{k \rightarrow \infty} |\varrho^k \nu_i(k) - \varrho^{k-1} \nu_i(k-1)| \\ &\leq \lim_{k \rightarrow \infty} [\varrho^k \frac{1}{\sqrt{3}\sigma} + \varrho^{k-1} \frac{1}{\sqrt{3}\sigma}] = 0. \end{aligned}$$

It means that the noises are decaying and converge to 0.

Next, we prove that the sum of all nodes' added noises equals to zero. Note that

$$\begin{aligned} \sum_{i=1}^n \sum_{k=0}^{\infty} \theta_i(k) &= \sum_{i=1}^n \theta_i(0) + \sum_{i=1}^n \theta_i(1) \\ &\quad + \sum_{i=1}^n \sum_{k=2}^{\infty} (\varrho^k \nu_i(k) - \varrho^{k-1} \nu_i(k-1)) \\ &= \sum_{i=1}^n \nu_i(0) + \sum_{i=1}^n (\varrho \nu_i(1) - \tilde{\nu}_i(0)) \\ &\quad + \sum_{i=1}^n (\varrho^\infty \nu_i(\infty) - \varrho^1 \nu_i(1)) \\ &= \sum_{i=1}^n \nu_i(0) + \sum_{i=1}^n \tilde{\nu}_i(0), \end{aligned}$$

where we have used the fact that  $\varrho^\infty \nu_i(\infty) = 0$ . Substituting (46) into the above equation yields that

$$\begin{aligned} \sum_{i=1}^n \sum_{k=0}^{\infty} \theta_i(k) &= \sum_{i=1}^n \nu_i(0) \\ &+ \sum_{i=1}^n \left[ \nu_i(0) - \sum_{j \in N_i} (F_{ij}(z_{ij}) - F_{ji}(z_{ji})) \right] \\ &= \sum_{i=1}^n \sum_{j \in N_i} [F_{ji}(z_{ji}) - F_{ij}(z_{ij})] = 0. \end{aligned}$$

Hence, we have completed the proof.  $\blacksquare$

**Theorem 5.2:** Suppose that the information set  $\mathcal{I}_i^1$  of node  $i$  is available to the other nodes and each node at least has two or more neighbors (i.e.,  $|N_i| \geq 2$  for all  $i \in V$ ). Then, the OPAC algorithm achieves  $(\epsilon, \delta)$ -data-privacy, where  $\delta = \frac{\epsilon}{\sqrt{3}\sigma}$  and  $\lim_{\epsilon \rightarrow 0} \delta = 0$ .

*Proof:* It has been known that when  $\mathcal{I}_i^1$  of node  $i$  is available to other nodes, its neighbor node  $j$  can obtain the real values of  $\theta_i(1), \theta_i(2), \dots, \theta_i(\infty)$ . Then, the value of  $\sum_{k=1}^{\infty} \theta_i(k)$  is released. Note that

$$\begin{aligned} \sum_{k=1}^{\infty} \theta_i(k) &= (\varrho^1 \nu_i(1) - \tilde{\nu}_i(0)) + \sum_{k=2}^{\infty} \theta_i(k) \\ &= (\varrho^1 \nu_i(1) - \tilde{\nu}_i(0)) + (\varrho^\infty \nu_i(\infty) - \varrho^1 \nu_i(1)) \\ &= \tilde{\nu}_i(0). \end{aligned}$$

It means that the value of  $\tilde{\nu}_i(0)$  is released and available to node  $j$ . From (46), one sees that  $\tilde{\nu}_i(0) \neq \theta_i(0)^1$  and

$$\tilde{\nu}_i(0) = \theta_i(0) - \sum_{j \in N_i} [F_{ij}(z_{ij}) - F_{ji}(z_{ji})]. \quad (48)$$

Since  $|N_i| \geq 2$  and  $F_{ij}$  is only known by node  $j$ , there exists  $F_{ij_o}(z_{ij_o}) - F_{j_o i}(z_{j_o i})$  for  $j_o \in N_i$  in the above equation is not known by node  $j$ . Meanwhile,  $F_{ij_o}(z_{ij_o}) - F_{j_o i}(z_{j_o i})$  could be any value in  $R$ , thus one infers that for any  $c \in [-\frac{1}{\sqrt{3}\sigma}, \frac{1}{\sqrt{3}\sigma}]$ ,

$$\Pr\{\theta_i(0) = c | \tilde{\nu}_i(0)\} = \Pr\{\theta_i(0) = c\}.$$

Hence, even if the value of  $\tilde{\nu}_i(0)$  is released, node  $j$  cannot increase the estimation accuracy of  $\theta_i(0)$  with (48). One thus concludes that based on the OPAC algorithm, the privacy compromise is avoided.

Note that  $x_i^+(0) = x_i(0) + \theta_i(0)$  is still available to node  $j$  for estimation and  $\theta_i(0) = \nu_i(0)$  follows uniform distribution. Then, with a similar analysis as Sec. IV-A, we have

$$\begin{aligned} \delta &= \max_{\hat{\theta}_i(0)} \int_{\hat{\theta}_i(0) - \epsilon}^{\hat{\theta}_i(0) + \epsilon} f_{\theta_i(0)}(y) dy \\ &= 2\epsilon \frac{1}{2\sqrt{3}\sigma} = \frac{\epsilon}{\sqrt{3}\sigma}, \end{aligned}$$

which completes the proof.  $\blacksquare$

If node  $i$  has only one neighbor node  $j$ , node  $j$  can infer the value of  $F_{ij}(z_{ij}) - F_{ji}(z_{ji})$ . Then, from (48), node  $j$  can obtain the value of  $\theta_i(0)$  and  $x_i(0)$  when  $\tilde{\nu}_i(0)$  is known.

<sup>1</sup>This is the main difference between OPAC and PPAC algorithm, and the main reason why OPAC can avoid privacy compromise.

**Remark 5.3:** From the above theorem, one sees that using OPAC algorithm, we have  $\delta = \frac{\epsilon}{\sqrt{3}\sigma}$ , which is the optimal privacy that can be achieved from solving problem (25). Furthermore,  $\delta = \frac{\epsilon}{\sqrt{3}\sigma}$  can be guaranteed by OPAC algorithm with  $\mathcal{I}_i^1(\infty)$ . Thus, OPAC algorithm can achieve a much higher  $(\epsilon, \delta)$ -data-privacy than PPAC.

## VI. PERFORMANCE EVALUATION

In this section, we conduct simulation to verify the obtained theoretical results and evaluate the performance of the proposed OPAC algorithm.

### A. Simulation Scenario

Consider the network with 50 nodes which are randomly deployed in a  $100\text{m} \times 100\text{m}$  area, and the maximum communication range of each node is 30m. We consider the normal distribution and uniform distribution of the added noises, respectively, where the mean and variance of them are set 0 and  $\sigma^2 = 1$ . We set  $\varrho = 0.9$ . The initial states of the nodes are randomly selected from  $[0, 10]$ . The function,  $d(t) = \max_{i \in V} |x_i(t) - \bar{x}|$ , is defined as the maximum difference between the nodes' states and the average value.

### B. Verification

Fig. 1(a) compares the convergence speed of the PPAC algorithm using normal and uniform distribution noises. It is observed that under the two different distributions, the PPAC algorithm has the same convergence speed. This justifies that the convergence speed only depends on the eigenvalues of the weighted matrix  $W$  and the value of  $\varrho$  as proved in [16].

Fig. 1(b) compares the  $(\epsilon, \delta)$ -data-privacy under  $\mathcal{I}_i^0$  with normal and uniform distribution noises. In simulation, we conduct 10,000 simulation runs. For each run, one node first generates a state  $\theta_i(0)$  randomly with the given distribution, and the other node generates 10,000 random numbers with the same distribution and use them as the estimation of  $\theta_i(0)$  (i.e.,  $\hat{\theta}_i(0)$ ). Then, one can get the probability of  $|\hat{\theta}_i(0) - \theta_i(0)| \leq \epsilon$  in each run, and we use the maximum probability among these in all runs simulation as the value of  $\delta$ . In theorem, we use (11) and (36) to calculate the value of  $\delta$  under two different distributions, respectively. Clearly, one can observe from Fig. 1(b) that uniform distribution is much better than normal distribution in the sense of  $(\epsilon, \delta)$ -data-privacy. It is also observed that the value of  $\delta$  in simulation and in theorem matches. Hence, this simulation verifies the theoretical results given in Sec. IV-A and B.

Fig. 1(c) compares the  $(\epsilon, \delta)$ -data-privacy under  $\mathcal{I}_i^1$  using normal and uniform distribution noises. The simulations here are conducted similarly as those in Fig. 1(b), except that when the iteration increases, the variance of the noises will be changed to  $\sigma^2 = \varrho^{2k}$  since (42) will be used for estimation at iteration  $k$ . We use (37) and (45) to calculate the value of  $\delta$  under the two different distributions, respectively, and the corresponding results are denoted as theoretical results. Both in simulation and theory, we set  $\epsilon = 0.2$ . As shown in Fig. 1(c), the maximum disclosure probability is increasing with



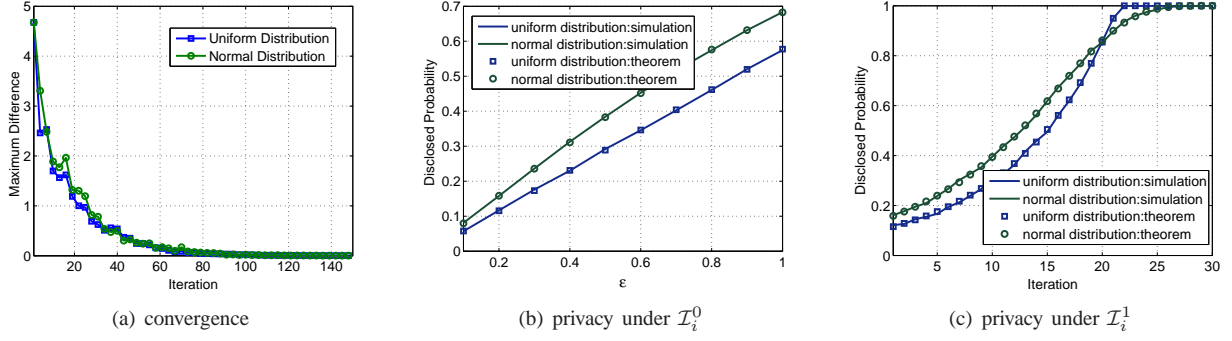


Fig. 1. The convergence and privacy comparison under different random noise distribution.

iteration and will converge to 1, i.e., the privacy is decaying with iteration and will eventually be lost. This simulation verifies the results obtained in Sec. IV-C.

### C. Evaluation

In this subsection, we will evaluate the performance of the OPAC algorithm. Using the same evaluation as the above section, the OPAC algorithm can guarantee the similar privacy as the blue line shown Fig. 1(b) under  $\mathcal{I}_i^1$ . This is because of that uniform distribution noise is used in OPAC and the secret function makes the later on ( $k \geq 1$ ) information cannot increase the disclosure probability. Therefore, the OPAC can guarantee much stronger privacy than PPAC, since it can achieve the same data-privacy under  $\mathcal{I}_i^1$  as PPAC under  $\mathcal{I}_i^0$ .

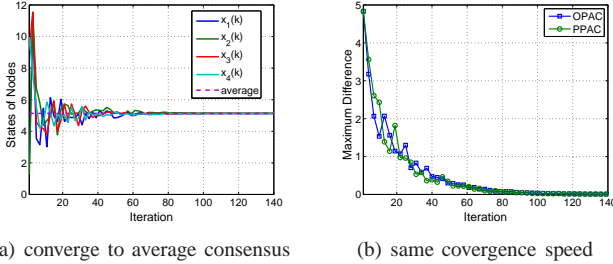


Fig. 2. The performance evaluation of the OPAC algorithm.

Then, we test the convergence of the OPAC algorithm. Set  $F_{ij} = \frac{i+j}{50}$ . As shown in Fig. 2(a), we find that the nodes' states will converge to the exact average with OPAC, which means that an exact average consensus can be achieved by the proposed algorithm. Fig. 2(b) compares the convergence speed of the OPAC and PPAC, it is found that they have the same convergence speed. It means that the added secret function will not affect the convergence speed.

## VII. CONCLUSIONS

In this paper, we investigated the privacy of an existing PPAC algorithm. We proposed a novel privacy definition, named  $(\epsilon, \delta)$ -data-privacy, to depict the relationship between the privacy and the estimation accuracy, so that the degree of the privacy can be well quantified. We proved that the PPAC algorithm achieves  $(\epsilon, \delta)$ -data-privacy, and obtained the closed-form expression of the relationship between  $\epsilon$  and  $\delta$ . We also proved that the uniform distribution noises can

guarantee a highest privacy when  $\epsilon$  is small enough. We revealed that the privacy will be lost when the information used in each consensus iteration is available to the other nodes. Then, to solve this problem and achieve highest  $(\epsilon, \delta)$ -data-privacy, we proposed OPAC algorithm, followed by the convergence and privacy analysis. Lastly, extensive simulation results are conducted to demonstrate the efficiency of the proposed algorithm.

## REFERENCES

- [1] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, 95(1): 215–233, 2007.
- [2] V. Blondel, J. M. Hendrickx, A. Olshevsky, and J. Tsitsiklis, "Convergence in multiagent coordination, consensus, and flocking," in *Proc. of IEEE CDC*, 2005.
- [3] W. Ren, B. Randal., and A. Ella, "Information consensus in multivehicle cooperative control: Collective group behavior through local interaction," *IEEE Control Systems Magazine*, 27(2): 71–82, 2007.
- [4] F. Pasqualetti, R. Carli, A. Bicchi, and F. Bullo, "Distributed estimation and detection under local information," in *Proc. of IFAC*, 2010.
- [5] G. Mateos, I. Schizas and G. Giannakis, "Distributed recursive least-squares for consensus-based in-network adaptive estimation," *IEEE Trans. Signal Processing*, 57(11): 4583–4588, 2009.
- [6] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. of ISIPSN*, 2005.
- [7] C. Zhao, J. He, P. Cheng and J. Chen, "Consensus-based energy management in smart grid with transmission losses and directed communication," *IEEE Trans. Smart Grid*, DOI: 10.1109/TSG.2015.2513772.
- [8] J. He, L. Duan, F. Hou, P. Cheng, and J. Chen, "Multi-period scheduling for wireless sensor networks: A distributed consensus approach," *IEEE Trans. Signal Processing*, 63(7): 1651–1663, 2015.
- [9] L. Schenato and F. Fiorentin, "Average timesynch: A consensus-based protocol for clock synchronization in wireless sensor networks," *Automatica*, 47(9): 1878–1886, 2011.
- [10] R. Carli, and S. Zampieri, "Network clock synchronization based on the second order linear consensus algorithm," *IEEE Trans Automat. Contr.*, 59(2): 409–422, 2014.
- [11] J. He, P. Cheng, L. Shi, and J. Chen, "Time synchronization in WSNs: A maximum value based consensus approach," *IEEE Trans Automat. Contr.*, 59(3): 660–674, 2014.
- [12] J. Le Ny and G. Pappas, "Differentially private filtering," *IEEE Trans Automat. Contr.*, 59(2): 341–354, 2014.
- [13] Z. Huang, S. Mitra, and G. Dullerud, "Differentially private iterative synchronous consensus," in *Proc. ACM workshop on Privacy in the electronic society*, 2012.
- [14] E. Nozari, P. Tallapragada, and J. Cortes, "Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design," *arXiv preprint arXiv:1512.09039*, 2016.
- [15] N. Maniara and C. Hadjicostis, "Privacy-preserving asymptotic average consensus," in *Proc. IEEE ECC*, 2013.
- [16] Y. Mo, and R. Murray, "Privacy preserving average consensus," in *Proc. IEEE CDC*, 2014.
- [17] Y. Mo, and R. Murray, "Privacy preserving average consensus," *IEEE Trans Automat Contr.*, accepted and to appear, 2016.

- [18] M. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, 69(345), 118–121, 1974.
- [19] J. Lin, A. S. Morse, and B. D. Anderson, "The multi-agent rendezvous problem," in *Proc. IEEE CDC*, 2003.
- [20] A. Olshevsky and J. Tsitsiklis, "Convergence speed in distributed consensus and averaging," *SIAM Review*, 53(4): 747–772, 2011.
- [21] C. Dwork, "Differential privacy," in *Automata, languages and programming*, Springer, 1-12, 2006.
- [22] J. He, L. Cai, P. Cheng, M. Xing, J. Pan and L. Shi. Private and accurate data aggregation against dishonest nodes. *Technical Report*, [http://web.uvic.ca/~jphe/privacy\\_aggregation.pdf](http://web.uvic.ca/~jphe/privacy_aggregation.pdf), 2016.